EarthCube Workbench

Background

EarthCube is a NSF Program from the Directorate for GeoSciences (GEO) and Directorate for Computer & Information Science & Engineering (CISE). The goal of the program is to "transform research in the academic geoscience community". This is achieved by the development of "cyberinfrastructure that is thoughtfully and systematically built to meet the current and future requirements of geoscientists."

To date, the EarthCube program has awarded 51 different projects. 25 of these are "building blocks", which are intended to be implemented by existing data facilities so that the capability is useful to a broad range of geoscientists. These building blocks are intended to become the 'enterprise architecture' for EarthCube .

https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=504780

https://www.nsf.gov/pubs/2016/nsf16514/nsf16514.htm

http://skc.isi.edu/wiki/earthcube/index.php/Building_Blocks

Earthcube "Workbench" conflation

It seems that there's some confusion around the idea of an Earthcube Workbench and how it relates to our NDS "Labs Workbench" project.

- On the All-Hands schedule are Workbench 101 and Science Driven Workbench sessions. It's unclear at this point what these refer to.
- An AGU presentation about an Interoperability Workbench
- 2016 EarthCube Architecture final report includes high-level outcome of the Geoscientist's Knowledge Workbench. (The term "workbench" is mentioned 44 times in this report)
    - "a flexible and collaborative environment where geoscientists and cyber infrastructure experts can discover, access and utilize an array of data and services that foster cross-domain advances in information and knowledge"

Building Blocks

The following projects were funded "building blocks" for EarthCube

https://www.earthcube.org/group-type/funded-projects-building-blocks?items_per_page=60

https://www.earthcube.org/tools-inventory

Name	Summary	Github	Workbench Notes
BCube	BCube crawler: Nutch, Solr, Elastic Map Reduce Broker based on GEOSS Websites links are broken https://nsidc.org/informatics/bcube/technology	https://github.com/b-cube/	Possible integration with Workbench, but unclear whether it fits based on available information.
CHORDS	http://chordsrt.com/ - defunct website http://portal.chordsrt.com/about is working Portal available to deploy on AWS via Docker	http://ncar.github.io/chords/ https://github.com/NCAR/chords	Seems like a reasonable candidate for Workbench, particularly since it's already deployed to Amazon via Docker.
CINERGI			May not be easily Dockerizable?
Cognitive Computer Infrastructure (GeoDeepDive)	Condor/Open Science Grid/DeepDive Input documents, OCR/NLP, output factor graph http://hazy.cs.wisc.edu/demo/geo/ defunct DeepDive is command line/API GeoDeepDive provides framework. There might have been a demo UI at some point		Seems to be primarily CLI and API/framework. Can be provided via workbench using either Cloud9 or Wetty interfaces. Probably not compelling to demonstrated.
CyberConnector	http://cube.csiss.gmu.edu/CyberConnector/web/index http://cube.csiss.gmu.edu/CyberConnector/web/demo Appears to be a catalog of web services and model transformation /integration service		It's not clear how this might fit workbench. Model integration/comparison services might make sense, but there's no clear deployable artifact. Primarily (at least according to the video) a centralized web service and Window application.

Integrated Data System			
Digital Crust		https://github.com/digitalcrust/	Not clear whether there is a software service here.
DisConnBB	https://groups.google.com/forum/#!forum/ec-disconbb http://nwisnfie-b.cloudapp.net/ defunct		No clear repository or service.
EarthSystemBridge	Interoperable modeling frameworks		Not software.
EarthCollab	Uses VIVO to model case studies to study collaboration		Not software
EC Data Discovery Hub	No information		No information
GeoDataSpace	Globus, but unclear how https://www.slideshare.net/TanuMalik/geodataspace-simplifying-data-management-tasks-with-globus		Not clear what this is
GeoLink	Reusable ontology design patterns http://www.geolink.org/ SPARQL endpoint		Primarily ontologies, but they seem to be publishable under SPARQL.
GeoSciCloud	Compare and contrast IT (cloud) infrastructure providers		Not software
Geosemantics	Developed at NCSA		Primarily a REST API
GeoTrust	Seems to be related to GeoDataSpace (also Tanu Malik)		Not clear what the software is.
GeoWS	Standard interface to geospatial data		A framework to present data conforming to a specific standard. Might be interesting to demonstrate, if we had an analysis client that could access the standard services.
Leveraging Semantics and Crowdsourcing			Not software
OceanLink	Now part of GeoLink		
ODSIP	Adds new functions to Data Access Protocol (DAP) framework to enable "Data as a Service"		
OntoSoft			In progress
Polar Data Insights	http://polar.usc.edu/ Banana for Solr, D3js It looks like a whole system for crawling polar data from a variety of sources, with a pipeline that uses DeepDive. They've also developed interfaces that sit on top of Solr, that could be interesting as re-usable components.		Possibly the Banana and D3.js services could be made available in Workbench.
Software Stewardship for the Geosciences			Not software
DRILSDOWN	Seems to be related to Jupyter notebooks?	https://github.com/Unidata/drilsdown	Appears to be plugin for RAMADDA repository? http://geodesystems.com/
Ensemble Toolkit			